

On the Connection Between Ritt Characteristic Sets and Buchberger-Gröbner Bases

Dongming Wang

School of Mathematics and Systems Science,
Beihang University, Beijing 100191, China

Centre National de la Recherche Scientifique,
75794 Paris cedex 16, France

Abstract

For any polynomial ideal \mathcal{I} , let the minimal triangular set contained in the reduced Buchberger-Gröbner basis of \mathcal{I} with respect to the purely lexicographical term order be called the W -characteristic set of \mathcal{I} . In this paper, we establish a strong connection between Ritt's characteristic sets and Buchberger's Gröbner bases of polynomial ideals by showing that the W -characteristic set \mathbb{C} of \mathcal{I} is a Ritt characteristic set of \mathcal{I} whenever \mathbb{C} is an ascending set, and a Ritt characteristic set of \mathcal{I} can always be computed from \mathbb{C} with simple pseudo-division when \mathbb{C} is regular. We also prove that under certain variable ordering, either the W -characteristic set of \mathcal{I} is normal, or irregularity occurs for the j th, but not the $(j+1)$ th, elimination ideal of \mathcal{I} for some j . In the latter case, we provide explicit pseudo-divisibility relations, which lead to nontrivial factorizations of certain polynomials in the Buchberger-Gröbner basis and thus reveal the structure of such polynomials. The pseudo-divisibility relations may be used to devise an algorithm to decompose arbitrary polynomial sets into normal triangular sets based on Buchberger-Gröbner bases computation.

Keywords: characteristic set, Gröbner basis, irregularity structure, polynomial ideal, triangular decomposition.

1 Introduction

In his classical works [8, 9], Ritt introduced the concept of characteristic sets for polynomial and differential polynomial ideals. This concept plays a central role in Ritt's theory of differential algebra. Ritt did not provide any effective way for the construction of characteristic sets even for non-prime polynomial ideals. Wu [11] developed Ritt's theory and method by working out efficient algorithms for the computation of characteristic sets of (differential) polynomial sets instead of (differential) polynomial ideals. On the other hand, in his Ph.D. thesis [2] Buchberger introduced the concept of Gröbner bases for polynomial ideals and proposed an algorithm for effective computation of Gröbner bases. Over the past three decades the theories and methods of characteristic sets and Gröbner bases have been studied extensively and independently by many researchers. They have become fundamental tools for computational commutative algebra and algebraic geometry (see, e.g., [1, 4, 5, 6]).

Characteristic sets and Gröbner bases are rather different in terms of structure and properties. The algorithms for their computation are distinguished by their elimination strategies (elimination of variables vs. elimination of terms) and reduction steps (pseudo-division vs. Buchberger reduction). It was not known what inherent connections may

exist between the characteristic sets and the Gröbner bases of arbitrary polynomial ideals. For any polynomial ideal \mathcal{I} , we call the minimal triangular set contained in the reduced Buchberger-Gröbner basis of \mathcal{I} with respect to the purely lexicographical term order the *W-characteristic set* of \mathcal{I} . This paper establishes, for the first time, a strong connection between Ritt's characteristic sets and Buchberger's Gröbner bases of polynomial ideals by showing that the W-characteristic set \mathbb{C} of \mathcal{I} is a Ritt characteristic set of \mathcal{I} whenever \mathbb{C} is an ascending set, and a Ritt characteristic set of \mathcal{I} can always be computed from \mathbb{C} with simple pseudo-division when \mathbb{C} is regular. We also prove that under certain variable ordering, either the W-characteristic set of \mathcal{I} is normal, or irregularity occurs for the j th, but not the $(j+1)$ th, elimination ideal of \mathcal{I} for some j . In the latter case, we provide explicit pseudo-divisibility relations, which lead to nontrivial factorizations of certain polynomials in the Buchberger-Gröbner basis and thus reveal the structure of such polynomials. It is not clear if the uniquely defined W-characteristic set of \mathcal{I} contains all what is needed for the construction of a Ritt characteristic set of \mathcal{I} , but it does appear to contain sufficient characteristic information about \mathcal{I} , in the abnormal case. The pseudo-divisibility relations may be used to devise an algorithm to decompose arbitrary polynomial sets into normal triangular sets based on Buchberger-Gröbner bases computation.

The strong connection between the Ritt characteristic set and the Buchberger-Gröbner basis of a polynomial ideal shown in this paper is simple, yet deep and surprising. It creates a route for the study of characteristic sets of polynomial ideals using the theory and method of Buchberger-Gröbner bases, and vice versa. Our work also illustrates how remarkable results in polynomial ideal theory can be obtained from the interplay of the two conceptually and operationally different methods. Further investigations on the W-characteristic sets of polynomial ideals are likely to help deepen our understanding of the structural properties of both characteristic sets and Buchberger-Gröbner bases.

2 Preliminaries

Let \mathcal{K} be a field and x_1, \dots, x_n be n variables with a fixed order $x_1 <_{\text{plex}} \dots <_{\text{plex}} x_n$. Denote by $\mathcal{K}[x_1, \dots, x_n]$ the ring of polynomials in x_1, \dots, x_n with coefficients in \mathcal{K} .

Let $F \in \mathcal{K}[x_1, \dots, x_n]$ be any nonzero polynomial. Denote by $\deg(F, x_k)$ the *degree* of F in x_k and by $\text{lc}(F, x_k)$ the *leading coefficient* of F with respect to x_k . Let $m = \deg(F, x_k)$ and G be any other polynomial of degree l in x_k . Pseudo-dividing G by F , considered as polynomials in x_k , one can obtain two polynomials Q and R in $\mathcal{K}[x_1, \dots, x_n]$ such that

$$I^q G = QF + R, \quad (1)$$

where

$$\begin{aligned} I &= \text{lc}(F, x_k), \quad q = \max(l - m + 1, 0), \\ \deg(R, x_k) &< m, \quad \deg(Q, x_k) = \max(l - m, -1). \end{aligned}$$

In case $m = 0$, $R = 0$ and $Q = G^l F$. The uniquely determined polynomials Q and R are called the *pseudo-quotient* and the *pseudo-remainder* of G with respect to F in x_k , denoted by $\text{pquo}(G, F, x_k)$ and $\text{prem}(G, F, x_k)$, respectively.

The polynomial G is said to be *R-reduced* with respect to F in x_k if $l < m$. When G is R-reduced with respect to F in x_k , $R = G$ and $Q = 0$.

Let $P \in \mathcal{K}[x_1, \dots, x_n] \setminus \mathcal{K}$ be any nonconstant polynomial. The biggest index p such that x_p actually occurs in P is called the *class* of P , denoted by $\text{cls}(P)$, and the variable x_p is called the *leading variable* of P , denoted by $\text{lv}(P)$. The class of any constant polynomial

in \mathcal{K} is defined to be 0. Let $\text{cls}(P) = p > 0$; then P can be written as $P = Ix_p^d + H$ with $I \in \mathcal{K}[x_1, \dots, x_{p-1}]$ and $\deg(H, x_k) < d = \deg(P, x_p)$. The leading coefficient I of P with respect to x_p is called the *initial* of P , denoted by $\text{ini}(P)$.

When a polynomial G is R-reduced with respect to P in $x_p = \text{lv}(P)$, we simply say that G is R-reduced with respect to P (without mentioning x_p).

2.1 Triangular sets and characteristic sets

To represent an ordered set, we enclose its elements using a pair of square brackets instead of braces.

Definition 2.1 A finite nonempty ordered set $[T_1, \dots, T_r]$ of nonconstant polynomials in $\mathcal{K}[x_1, \dots, x_n]$ is called a *triangular set* if $0 < \text{cls}(T_1) < \dots < \text{cls}(T_r)$.

A finite nonempty ordered set $\mathbb{A} = [A_1, \dots, A_r]$ of nonzero polynomials in $\mathcal{K}[x_1, \dots, x_n]$ is called an *ascending set* if either $r = 1$ and $A_1 \in \mathcal{K}$, or \mathbb{A} is a triangular set and A_j is R-reduced with respect to A_i for every pair $i < j$ and $j = 2, \dots, r$.

Let $\mathbb{T} = [T_1, \dots, T_r]$ be any triangular set and P be any polynomial in $\mathcal{K}[x_1, \dots, x_n]$. P is said to be *R-reduced* with respect to \mathbb{T} if P is R-reduced with respect to every $T_i \in \mathbb{T}$, i.e., $\deg(P, \text{lv}(T_i)) < \deg(T_i, \text{lv}(T_i))$ for all $1 \leq i \leq r$.

The polynomial

$$R = \text{prem}(\dots \text{prem}(P, T_r, \text{lv}(T_r)), \dots, T_1, \text{lv}(T_1)),$$

denoted simply by $\text{prem}(P, \mathbb{T})$, is called the *pseudo-remainder* of P with respect to \mathbb{T} . From (1), one can easily derive the following *pseudo-remainder formula*

$$I_1^{q_1} \dots I_r^{q_r} P = Q_1 T_1 + \dots + Q_r T_r + R, \quad (2)$$

where each q_i is a nonnegative integer and

$$I_i = \text{ini}(T_i), \quad Q_i \in \mathcal{K}[x_1, \dots, x_n], \quad i = 1, \dots, r.$$

Definition 2.2 For any two nonzero polynomials F and G in $\mathcal{K}[x_1, \dots, x_n]$, F is said to have *lower rank* than G , denoted as $F \prec G$ or $G \succ F$, if either $\text{cls}(F) < \text{cls}(G)$, or $\text{cls}(F) = \text{cls}(G) > 0$ and $\deg(F, \text{lv}(F)) < \deg(G, \text{lv}(G))$. In this case, G is said to have *higher rank* than F .

If neither $F \prec G$ nor $G \prec F$, then F and G are said to have the *same rank*, denoted as $F \sim G$.

Definition 2.3 For any two ascending sets

$$\mathbb{A} = [A_1, \dots, A_r], \quad \mathbb{A}' = [A'_1, \dots, A'_{r'}],$$

\mathbb{A} is said to have *higher rank* than \mathbb{A}' , denoted as $\mathbb{A} \succ \mathbb{A}'$ or $\mathbb{A}' \prec \mathbb{A}$, if one of the following conditions holds:

(a) there exists an integer $j \leq \min(r, r')$ such that

$$A_1 \sim A'_1, \dots, A_{j-1} \sim A'_{j-1}, \quad \text{while } A_j \succ A'_j;$$

(b) $r' > r$ and $A_1 \sim A'_1, \dots, A_r \sim A'_r$.

In this case, \mathbb{A}' is said to have *lower rank* than \mathbb{A} . If neither $\mathbb{A} \prec \mathbb{A}'$ nor $\mathbb{A}' \prec \mathbb{A}$, then \mathbb{A} and \mathbb{A}' are said to have the *same rank*, denoted as $\mathbb{A} \sim \mathbb{A}'$. In this case,

$$r = r', \quad \text{and} \quad A_1 \sim A'_1, \dots, A_r \sim A'_r.$$

For any polynomial set $\mathbb{P} = \{P_1, \dots, P_s\}$ in $\mathcal{K}[x_1, \dots, x_n]$, denote by $\langle \mathbb{P} \rangle$ or $\langle P_1, \dots, P_s \rangle$ the ideal generated by the polynomials P_1, \dots, P_s in $\mathcal{K}[x_1, \dots, x_n]$ and by $\text{Zero}(\mathbb{P})$ the set of all common zeros (in some extension field of \mathcal{K}) of P_1, \dots, P_s .

Definition 2.4 Let \mathbb{P} be any finite nonempty polynomial set in $\mathcal{K}[x_1, \dots, x_n]$. With respect to the rank \succ , any minimal ascending set contained in $\langle \mathbb{P} \rangle$ is called a *Ritt characteristic set* of $\langle \mathbb{P} \rangle$.

Lemma 2.1 For any finite nonempty polynomial set \mathbb{P} in $\mathcal{K}[x_1, \dots, x_n]$, an ascending set \mathbb{A} contained in $\langle \mathbb{P} \rangle$ is a Ritt characteristic set of $\langle \mathbb{P} \rangle$ if and only if $\text{prem}(P, \mathbb{A}) \equiv 0$ for all $P \in \langle \mathbb{P} \rangle$.

Proof. See [7, Theorem 5.3.3]. \square

Definition 2.5 For any triangular set $\mathbb{T} = [T_1, \dots, T_r]$ in $\mathcal{K}[x_1, \dots, x_n]$ with $I_i = \text{ini}(T_i)$ for $1 \leq i \leq r$, the *saturated ideal* of \mathbb{T} is defined to be

$$\text{sat}(\mathbb{T}) = \{F \in \mathcal{K}[x_1, \dots, x_n] \mid (I_1 \cdots I_r)^q F \in \langle \mathbb{T} \rangle \text{ for some integer } q > 0\}.$$

2.2 Regular sets and normal triangular sets

For two polynomials F and G in $\mathcal{K}[x_1, \dots, x_n]$, denote by $\text{res}(F, G, x_k)$ the *resultant* of F and G with respect to x_k . For any polynomial F and triangular set $\mathbb{T} = [T_1, \dots, T_r]$, define $\text{res}(F, \mathbb{T}) = \text{res}(\cdots \text{res}(F, T_r, \text{lv}(T_r)), \dots, T_1, \text{lv}(T_1))$. It is easy to see that there exist polynomials $A, B_1, \dots, B_r \in \mathcal{K}[x_1, \dots, x_n]$ such that

$$\text{res}(F, \mathbb{T}) = AF + B_1 T_1 + \cdots + B_r T_r \quad (3)$$

(cf. [7, Lemmas 7.2.1 and 7.2.2] and [10, Lemma 4.3.2]), so $\text{res}(F, \mathbb{T}) \in \langle F, T_1, \dots, T_r \rangle$.

Definition 2.6 Let $\mathbb{T} = [T_1, \dots, T_r]$ be any triangular set in $\mathcal{K}[x_1, \dots, x_n]$. \mathbb{T} is said to be *regular* or called a *regular set* if $\text{res}(\text{ini}(T_j), [T_1, \dots, T_{j-1}]) \not\equiv 0$ for all $j = 2, \dots, r$. \mathbb{T} is said to be *normal* if $\deg(\text{ini}(T_j), \text{lv}(T_i)) = 0$ for all $i < j$ and $j = 2, \dots, r$.

Regular sets, also known as *regular chains*, have other equivalent definitions (see, e.g., [10, p.114]). Here they are defined by means of resultants for the convenience of proof in Section 3. A triangular set \mathbb{T} as in Definition 2.6 is regular if and only if the image of $\text{ini}(T_i)$ is neither zero nor a zero divisor in the quotient ring $\mathcal{K}[x_1, \dots, x_n]/\text{sat}([T_1, \dots, T_{i-1}])$ for all $i = 2, \dots, r$.

Proposition 2.2 Every normal triangular set in $\mathcal{K}[x_1, \dots, x_n]$ is regular.

Proof. It is obvious. \square

Proposition 2.3 Let \mathbb{T} be any triangular set in $\mathcal{K}[x_1, \dots, x_n]$. Then \mathbb{T} is regular if and only if $\text{prem}(P, \mathbb{T}) \equiv 0$ for all $P \in \text{sat}(\mathbb{T})$.

Proof. See [10, Theorem 6.2.4] and references therein. \square

Lemma 2.4 *Let $\mathbb{T} = [T_1, \dots, T_r]$ with $p_r = \text{cls}(T_r) < n$ be any regular set and $P = P_d x_m^d + \dots + P_1 x_m + P_0$ with $m > p_r$ and $d = \deg(P, x_m) > 0$ be any polynomial in $\mathcal{K}[x_1, \dots, x_n]$. Then $\text{prem}(P, \mathbb{T}) \equiv 0$ if and only if $\text{prem}(P_j, \mathbb{T}) \equiv 0$ for all $0 \leq j \leq d$.*

Proof. Let $I_i = \text{ini}(T_i)$ for $1 \leq i \leq r$. Suppose that $\text{prem}(P, \mathbb{T}) \equiv 0$. Then according to (2) there exist a power product K of I_1, \dots, I_r (i.e., $K = I_1^{q_1} \dots I_r^{q_r}$ for some nonnegative integers q_1, \dots, q_r) and polynomials Q_1, \dots, Q_r in $\mathcal{K}[x_1, \dots, x_n]$ such that $KP = Q_1 T_1 + \dots + Q_r T_r$. It follows that

$$KP_j = \text{coef}(Q_1 T_1 + \dots + Q_r T_r, x_m^j) = \text{coef}(Q_1, x_m^j) T_1 + \dots + \text{coef}(Q_r, x_m^j) T_r \in \langle \mathbb{T} \rangle$$

for $0 \leq j \leq d$, where $\text{coef}(F, x_m^j)$ denotes the coefficient of F in x_m^j . Therefore, $P_j \in \text{sat}(\mathbb{T})$ by definition. As \mathbb{T} is regular, we have $\text{prem}(P_j, \mathbb{T}) \equiv 0$ according to Proposition 2.3.

To show the other direction, suppose that $\text{prem}(P_j, \mathbb{T}) \equiv 0$ for all $0 \leq j \leq d$. Then $P_j \in \text{sat}(\mathbb{T})$ for all j . This implies that $P \in \text{sat}(\mathbb{T})$. Since \mathbb{T} is regular, we have $\text{prem}(P, \mathbb{T}) \equiv 0$, again by Proposition 2.3. The proof is complete. \square

Corollary 2.5 *Let \mathbb{T} be any regular set and P and F be any two polynomials in $\mathcal{K}[x_1, \dots, x_n]$. If $\text{prem}(P, \mathbb{T}) = \text{prem}(F, \mathbb{T}) \equiv 0$, then $\text{prem}(P + F, \mathbb{T}) \equiv 0$.*

Proof. Let y be a variable not occurring in P, F , and \mathbb{T} . By Lemma 2.4, $\text{prem}(P, \mathbb{T}) = \text{prem}(F, \mathbb{T}) \equiv 0$ implies that $\text{prem}(Py + F, \mathbb{T}) \equiv 0$. The corollary is proved by taking $y = 1$. \square

Lemma 2.4 and Corollary 2.5 do not hold when \mathbb{T} is not regular. This can be seen from the simple example, where $\mathbb{T} = [x_1^2, x_1 x_2]$, $P = x_3 - x_2^2$, and $F = x_2^2$; it is easy to see that $\text{prem}(P, \mathbb{T}) = \text{prem}(F, \mathbb{T}) \equiv 0$, but $\text{prem}(\text{lc}(P, x_3), \mathbb{T}) \not\equiv 0$ and $\text{prem}(P + F, \mathbb{T}) \not\equiv 0$.

2.3 Buchberger-Gröbner bases

Two distinct monomials $x_1^{i_1} \dots x_n^{i_n}$ and $x_1^{j_1} \dots x_n^{j_n}$ in x_1, \dots, x_n are ordered as

$$x_1^{i_1} \dots x_n^{i_n} <_{\text{plex}} x_1^{j_1} \dots x_n^{j_n} \quad \text{or} \quad x_1^{j_1} \dots x_n^{j_n} >_{\text{plex}} x_1^{i_1} \dots x_n^{i_n},$$

if there exists an integer k ($1 \leq k \leq n$) such that

$$i_n = j_n, \dots, i_{k+1} = j_{k+1} \quad \text{while} \quad i_k < j_k.$$

Under $<_{\text{plex}}$, all the monomials in x_1, \dots, x_n are ordered, and so are the terms of any nonzero polynomial in $\mathcal{K}[x_1, \dots, x_n]$. We call $<_{\text{plex}}$ the *purely lexicographical order* (plex) of monomials or terms.

Any nonzero polynomial P in $\mathcal{K}[x_1, \dots, x_n]$ can be written in the form

$$P = \sum_{l=1}^t a_l x_1^{i_{l1}} \dots x_n^{i_{ln}}$$

with

$$a_1 \neq 0, \dots, a_t \neq 0, \quad a_i \in \mathcal{K}, \\ x_1^{i_{11}} \dots x_n^{i_{1n}} >_{\text{plex}} \dots >_{\text{plex}} x_1^{i_{t1}} \dots x_n^{i_{tn}}.$$

We call $x_1^{i_{11}} \cdots x_n^{i_{1n}}$ the *leading monomial*, $a_1 x_1^{i_{11}} \cdots x_n^{i_{1n}}$ the *leading term*, a_1 the *leading coefficient* of P , and a_l the *coefficient* of P in $x_1^{i_{l1}} \cdots x_n^{i_{ln}}$, denoted by $\text{lm}(P)$, $\text{lt}(P)$, $\text{lc}(P)$, and $\text{coef}(P, x_1^{i_{l1}} \cdots x_n^{i_{ln}})$, respectively. If Q is another nonzero polynomial in $\mathcal{K}[x_1, \dots, x_n]$, we order P and Q as $P <_{\text{plex}} Q$ or $Q >_{\text{plex}} P$ if $\text{lm}(P) <_{\text{plex}} \text{lm}(Q)$.

Definition 2.7 Let \mathbb{P} be any finite nonempty polynomial set and G be any polynomial in $\mathcal{K}[x_1, \dots, x_n]$. G is said to be *B-reducible* with respect to \mathbb{P} if there exist a polynomial $P \in \mathbb{P}$ and a monomial λ such that $\text{coef}(G, \lambda \text{lm}(P)) \neq 0$. If no such P and λ exist, G is said to be *B-reduced* or in *normal form* with respect to \mathbb{P} .

If G is B-reducible with respect to \mathbb{P} , then one can find a polynomial $P \in \mathbb{P}$ with the monomial $\lambda \text{lm}(P)$ maximal (with respect to the term order $<_{\text{plex}}$) such that

$$G = b \lambda P + H,$$

where

$$b = \frac{\text{coef}(G, \lambda \text{lm}(P))}{\text{lc}(P)}.$$

If H is B-reducible with respect to \mathbb{P} , then one can reduce H to another polynomial in the same way by choosing P, b , and λ . Such a process will terminate after a finite number of reduction steps. The finally obtained polynomial N will be B-reduced with respect to \mathbb{P} . In this case, one gets a formula of the form

$$G = Q_1 P_1 + \cdots + Q_s P_s + N,$$

in which $P_j \in \mathbb{P}$, $Q_j, N \in \mathcal{K}[x_1, \dots, x_n]$ and N is B-reduced with respect to \mathbb{P} . The polynomial N is called the *normal form* of G with respect to \mathbb{P} and denoted by $\text{nform}(G, \mathbb{P})$.

Definition 2.8 Let \mathbb{P} be an arbitrary finite and nonempty set of nonzero polynomials in $\mathcal{K}[x_1, \dots, x_n]$. A polynomial set \mathbb{G} in $\mathcal{K}[x_1, \dots, x_n]$ is called the *reduced Buchberger-Gröbner basis* of $\langle \mathbb{P} \rangle$ or \mathbb{P} with respect to the plex term order determined by $x_1 <_{\text{plex}} \cdots <_{\text{plex}} x_n$, if

- (a) for all $P \in \mathcal{K}[x_1, \dots, x_n]$, $P \in \langle \mathbb{P} \rangle$ if and only if $\text{nform}(P, \mathbb{G}) = 0$;
- (b) every polynomial $G \in \mathbb{G}$ is monic and B-reduced with respect to $\mathbb{G} \setminus \{G\}$;
- (c) $\langle \mathbb{G} \rangle = \langle \mathbb{P} \rangle$.

The reduced Buchberger-Gröbner basis of $\langle \mathbb{P} \rangle$ is unique and can be computed from \mathbb{P} by using Buchberger's algorithm [3]. What is called Buchberger-Gröbner basis here was named Gröbner basis by Buchberger after his Ph.D. advisor Wolfgang Gröbner. The author feels that the basis should be named more appropriately also after Bruno Buchberger for his outstanding contributions to the development of the theory and method of Gröbner bases.

3 Main Results

For any (finite or infinite) polynomial set $\mathbb{F} \subset \mathcal{K}[x_1, \dots, x_n]$ and $0 \leq j \leq k \leq n$, let $\mathbb{F}^{(j, \dots, k)}$ stand for $(\mathbb{F} \cap \mathcal{K}[x_1, \dots, x_k]) \setminus (\mathbb{F} \cap \mathcal{K}[x_1, \dots, x_{j-1}])$. When $j = k$, $\mathbb{F}^{(j, \dots, k)}$ is written as $\mathbb{F}^{(k)}$.

Definition 3.1 Let \mathbb{P} be an arbitrary finite and nonempty set of nonzero polynomials in $\mathcal{K}[x_1, \dots, x_n]$ and \mathbb{G} be the reduced Buchberger-Gröbner basis of \mathbb{P} with respect to the plex term order determined by $x_1 <_{\text{plex}} \dots <_{\text{plex}} x_n$. The set

$$\bigcup_{i=0}^n \left\{ G \mid G \in \mathbb{G}^{(i)}; G' >_{\text{plex}} G, \text{ for all } G' \in \mathbb{G}^{(i)} \setminus \{G\} \right\}$$

of polynomials, ordered by $<_{\text{plex}}$, is called the *W-characteristic set* of $\langle \mathbb{P} \rangle$.

The above-defined W-characteristic set \mathbb{C} is obviously a triangular set, but it is not necessarily an ascending set. \mathbb{C} is minimal in the sense that (i) each element C of \mathbb{C} has the lowest plex order among all those polynomials in the Buchberger-Gröbner basis \mathbb{G} which have the same leading variable as C and (ii) the number of elements in \mathbb{C} is the maximum possible. Owing to the uniqueness of the reduced Buchberger-Gröbner basis, the W-characteristic set of a polynomial ideal is uniquely defined. We will see that W-characteristic sets can be effectively used to bridge Ritt characteristic sets and Buchberger-Gröbner bases.

3.1 Construction of Ritt characteristic sets

In what follows, let \mathbb{P} be an arbitrary finite and nonempty set of nonzero polynomials in $\mathcal{K}[x_1, \dots, x_n]$, let \mathbb{G} be the reduced Buchberger-Gröbner basis of \mathbb{P} with respect to the plex term order determined by $x_1 <_{\text{plex}} \dots <_{\text{plex}} x_n$, and assume that $\mathbb{G} \neq [1]$, so $\mathbb{C} \neq [1]$, whenever needed. The following proposition shows that the W-characteristic set of $\langle \mathbb{P} \rangle$ possesses the main properties that a Ritt characteristic set of $\langle \mathbb{P} \rangle$ has.

Proposition 3.1 Let $\mathbb{C} = [C_1, \dots, C_r]$ be the W-characteristic set of $\langle \mathbb{P} \rangle$ with $I_i = \text{ini}(C_i)$ for $1 \leq i \leq r$. Then:

- (a) $\text{prem}(P, \mathbb{C}) \equiv 0$ for all $P \in \langle \mathbb{P} \rangle$;
- (b) $\langle \mathbb{C} \rangle \subset \langle \mathbb{P} \rangle \subset \text{sat}(\mathbb{C})$;
- (c) $\text{Zero}(\mathbb{C}) \setminus \text{Zero}(\{I_1 \cdots I_r\}) \subset \text{Zero}(\mathbb{P}) \subset \text{Zero}(\mathbb{C})$.

Proof. (a) Let $p_i = \text{cls}(C_i)$ for $1 \leq i \leq r$. Then $\mathbb{G} = \mathbb{G}^{(p_1)} \cup \dots \cup \mathbb{G}^{(p_r)}$ and for all $C \in \mathbb{G}^{(p_i)} \setminus \{C_i\}$, $C >_{\text{plex}} C_i$ and thus $\deg(\text{lm}(C), x_{p_i}) = \deg(C, x_{p_i}) \geq \deg(C_i, x_{p_i})$. Hence, for any $P \in \langle \mathbb{P} \rangle$, $R = \text{prem}(P, \mathbb{C})$ is B-reduced with respect to \mathbb{G} . Therefore, $R \equiv 0$ and (a) is proved.

(b) Note that $\mathbb{C} \subset \mathbb{G} \subset \langle \mathbb{P} \rangle$, so $\langle \mathbb{C} \rangle \subset \langle \mathbb{P} \rangle$. For any $P \in \langle \mathbb{P} \rangle$, by (a) and the pseudo-remainder formula for $\text{prem}(P, \mathbb{C}) = 0$ there exist nonnegative integers q_1, \dots, q_r such that $I_1^{q_1} \cdots I_r^{q_r} P \in \langle \mathbb{C} \rangle$. It follows from the definition of saturated ideals that $P \in \text{sat}(\mathbb{C})$. Therefore, (b) is proved.

(c) $\text{Zero}(\mathbb{P}) \subset \text{Zero}(\mathbb{C})$ follows from the first \subset relation in (b). Consider any zero $\mathbf{a} \in \text{Zero}(\mathbb{C}) \setminus \text{Zero}(\{I_1 \cdots I_r\})$ and let P be any polynomial in \mathbb{P} . Then $C_i(\mathbf{a}) = 0$ and $I_i(\mathbf{a}) \neq 0$ for $1 \leq i \leq r$. Plunging \mathbf{a} into the pseudo-remainder formula for $\text{prem}(P, \mathbb{C}) = 0$, we see that $P(\mathbf{a}) = 0$. Therefore, $\mathbf{a} \in \text{Zero}(\mathbb{P})$ and (c) is proved. \square

Proposition 3.2 Let \mathbb{C} be the W-characteristic set of $\langle \mathbb{P} \rangle$. Then for every i ($0 \leq i \leq n$), $\mathbb{C}^{(0, \dots, i)}$ is the W-characteristic set of the $(n-i)$ th elimination ideal $\langle \mathbb{P} \rangle^{(0, \dots, i)}$ of $\langle \mathbb{P} \rangle$.

Proof. By the elimination theorem of Buchberger-Gröbner bases (see, e.g., [4, Ch. 3, Theorem 2]), $\mathbb{G}^{(0, \dots, i)}$ is the reduced plex Buchberger-Gröbner basis of $\langle \mathbb{P} \rangle^{(0, \dots, i)}$. Hence the W -characteristic set of $\langle \mathbb{P} \rangle^{(0, \dots, i)}$ is identical to $\mathbb{C}^{(0, \dots, i)}$. \square

Theorem 3.3 *Let \mathbb{C} be the W -characteristic set of $\langle \mathbb{P} \rangle$. If \mathbb{C} is an ascending set, then \mathbb{C} is a Ritt characteristic set of $\langle \mathbb{P} \rangle$.*

Proof. By Proposition 3.1 (a), $\text{prem}(P, \mathbb{C}) \equiv 0$ for all $P \in \langle \mathbb{P} \rangle$. The theorem follows from Lemma 2.1. \square

Theorem 3.4 *Let $\mathbb{C} = [C_1, \dots, C_r]$ be the W -characteristic set of $\langle \mathbb{P} \rangle$. If \mathbb{C} is regular, then*

$$\mathbb{C}^* = [C_1, \text{prem}(C_2, [C_1]), \dots, \text{prem}(C_r, [C_1, \dots, C_{r-1}])]$$

is a Ritt characteristic set of $\langle \mathbb{P} \rangle$, where \mathbb{C}^ is also regular.*

Proof. Let $\mathbb{C}_i = [C_1, \dots, C_i]$, $x_{p_i} = \text{lv}(C_i)$, $I_i = \text{ini}(C_i)$, and $\mathbb{C}_i^* = [C_1^*, \dots, C_i^*]$ for $1 \leq i \leq r$, where $C_1^* = C_1$ and $C_i^* = \text{prem}(C_i, \mathbb{C}_{i-1})$ for $2 \leq i \leq r$. Note first that $C_i^* \in \langle \mathbb{C} \rangle$, so $\langle \mathbb{C}^* \rangle \subset \langle \mathbb{C} \rangle$. Let J be any power product of $I_1^* = \text{lc}(C_1^*, x_{p_1}), \dots, I_r^* = \text{lc}(C_r^*, x_{p_r})$. Observe from the pseudo-remainder formula for $C_i^* = \text{prem}(C_i, \mathbb{C}_{i-1})$ that for each I_i^* , there exist a power product J_i of I_1, \dots, I_i and polynomials $Q_{i,1}, \dots, Q_{i,i-1}$ in $\mathcal{K}[x_1, \dots, x_n]$ such that

$$J_i - (Q_{i,1}C_1 + \dots + Q_{i,i-1}C_{i-1}) = I_i^*.$$

Multiplying the two sides of such equalities up to certain powers, one sees that J is equal to a power product H of J_1, \dots, J_r plus a linear combination $F = Q_1C_1 + \dots + Q_{r-1}C_{r-1}$ for some polynomials $Q_i \in \mathcal{K}[x_1, \dots, x_n]$, i.e., $J = H + F$, so $H - J \in \langle \mathbb{C} \rangle$. As H is also a power product of I_1, \dots, I_r and \mathbb{C} is regular, $N = \text{res}(H, \mathbb{C}) \neq 0$ and N does not involve x_{p_1}, \dots, x_{p_r} . According to (3) there exists a polynomial A in $\mathcal{K}[x_1, \dots, x_n]$ such that $N - AH \in \langle \mathbb{C} \rangle$. This implies that $N - AJ \in \langle \mathbb{C} \rangle$. Therefore, $J \neq 0$; for otherwise, $N = \text{prem}(N, \mathbb{C}) = \text{prem}(N - AJ, \mathbb{C}) \equiv 0$ (by Proposition 2.3) leads to contradiction. In particular, we have $I_i^* \neq 0$ for all i . Since $\deg(C_i^*, x_{p_i})$ cannot be greater than $\deg(C_i, x_{p_i})$, they must be equal. This shows that $\text{lv}(C_i^*) = x_{p_i}$ and $\text{ini}(C_i^*) = I_i^*$ for $1 \leq i \leq r$. Consequently, C_i^* is R -reduced with respect to \mathbb{C}_{i-1}^* , as it is so with respect to \mathbb{C}_{i-1} , for $2 \leq i \leq r$; thereby \mathbb{C}^* is an ascending set.

For any polynomial $P \in \text{sat}(\mathbb{C}^*)$, there exists a power product J of I_1^*, \dots, I_r^* such that $JP \in \langle \mathbb{C}^* \rangle \subset \langle \mathbb{C} \rangle$. According to the above reasoning, we have $N - AJ \in \langle \mathbb{C} \rangle$. It follows that $NP \in \langle \mathbb{C} \rangle$. On the other hand, let $R^* = \text{prem}(P, \mathbb{C}^*)$. Then there exists a power product K of I_1^*, \dots, I_r^* such that $KP - R^* \in \langle \mathbb{C}^* \rangle \subset \langle \mathbb{C} \rangle$. It follows that $NR^* \in \langle \mathbb{C} \rangle$. Obviously, $\deg(R^*, x_{p_i}) < \deg(C_i^*, x_{p_i}) = \deg(C_i, x_{p_i})$ for all i . Hence $NR^* = \text{prem}(NR^*, \mathbb{C}) \equiv 0$. Therefore, $R^* \equiv 0$ and thus \mathbb{C}^* is regular by Proposition 2.3.

Next we want to show that $\langle \mathbb{C} \rangle \subset \text{sat}(\mathbb{C}^*)$. For this purpose, assume by induction that $C_{i-1} \in \text{sat}(\mathbb{C}_{i-1}^*)$; the case for C_1 is trivial. Then there exists a power product K_i of I_1, \dots, I_{i-1} such that $K_iC_i - C_i^* \in \langle \mathbb{C}_{i-1} \rangle \subset \text{sat}(\mathbb{C}_{i-1}^*)$. Since \mathbb{C} is regular, $R_i = \text{res}(K_i, \mathbb{C}_{i-1}) \neq 0$ and R_i does not involve x_{p_1}, \dots, x_{p_r} for each i . According to (3) there exists a polynomial A_i in $\mathcal{K}[x_1, \dots, x_n]$ such that $R_i - A_iK_i \in \langle \mathbb{C}_{i-1} \rangle \subset \text{sat}(\mathbb{C}_{i-1}^*)$. This implies that $R_iC_i \in \text{sat}(\mathbb{C}_i^*)$. It follows from Proposition 2.3 that

$$R_i \text{prem}(C_i, \mathbb{C}_i^*) = \text{prem}(R_iC_i, \mathbb{C}_i^*) \equiv 0.$$

Hence $\text{prem}(C_i, \mathbb{C}_i^*) \equiv 0$ and $C_i \in \text{sat}(\mathbb{C}_i^*)$. Therefore, $\langle \mathbb{C} \rangle \subset \text{sat}(\mathbb{C}^*)$.

Now consider any polynomial $P \in \langle \mathbb{P} \rangle$. Clearly, $R = \text{prem}(P, \mathbb{C})$ is B-reduced with respect to the Buchberger-Gröbner basis \mathbb{G} of \mathbb{P} . This implies that $R \equiv 0$ and $P \in \text{sat}(\mathbb{C})$. Hence there exists a power product L of I_1, \dots, I_r such that $LP \in \langle \mathbb{C} \rangle$. Since $M = \text{res}(L, \mathbb{C}) \neq 0$, there exists a polynomial B in $\mathcal{K}[x_1, \dots, x_n]$ such that $M - BL \in \langle \mathbb{C} \rangle$. It follows that $MP \in \langle \mathbb{C} \rangle \subset \text{sat}(\mathbb{C}^*)$. As M does not involve x_{p_1}, \dots, x_{p_r} , $M \text{prem}(P, \mathbb{C}^*) = \text{prem}(MP, \mathbb{C}^*) \equiv 0$ according to Proposition 2.3; therefore $\text{prem}(P, \mathbb{C}^*) \equiv 0$. By Lemma 2.1, \mathbb{C}^* is a Ritt characteristic set of $\langle \mathbb{P} \rangle$. \square

Corollary 3.5 *Let \mathbb{C} be the W-characteristic set of $\langle \mathbb{P} \rangle$. For any $0 \leq i \leq n$, if $\mathbb{C}^{(0, \dots, i)} = [C_1, \dots, C_k]$ is regular, then*

$$[C_1, \text{prem}(C_2, [C_1]), \dots, \text{prem}(C_k, [C_1, \dots, C_{k-1}])]$$

is a Ritt characteristic set of the elimination ideal $\langle \mathbb{P} \rangle^{(0, \dots, i)}$.

Proof. It follows from Proposition 3.2 and Theorem 3.4. \square

3.2 Structure of Buchberger-Gröbner bases

To explore the structural properties of the W-characteristic set \mathbb{C} of $\langle \mathbb{P} \rangle$, we assume from now on that x_1, \dots, x_n are properly ordered such that the leading variables x_{p_i} of the polynomials in \mathbb{C} are greater than all the other free variables, called *parameters*. Let $y_1 = x_{p_1}, \dots, y_r = x_{p_r}$ and u_1, \dots, u_m be all the parameters, where $m + r = n$. Then the assumed variable order is $u_1 <_{\text{plex}} \dots <_{\text{plex}} u_m <_{\text{plex}} y_1 <_{\text{plex}} \dots <_{\text{plex}} y_r$.

Theorem 3.6 *Let $[C_1, \dots, C_r]$ be the W-characteristic set of $\langle \mathbb{P} \rangle$. For any $1 \leq k < r$, if $\mathbb{C}_k = [C_1, \dots, C_k]$ is normal and $I_{k+1} = \text{ini}(C_{k+1})$, with $\text{lv}(I_{k+1}) = y_l$, is not R-reduced with respect to C_l , then*

$$\text{prem}(I_{k+1}, \mathbb{C}_k) \equiv 0 \quad \text{and} \quad \text{prem}(C_{k+1}, \mathbb{C}_k) \equiv 0.$$

Proof. Let $y_i = \text{lv}(C_i)$ and $I_i = \text{ini}(C_i)$ for $1 \leq i \leq r$. Suppose that I_{k+1} is not R-reduced with respect to C_l and let $R = \text{prem}(I_{k+1}, \mathbb{C}_k)$. Then according to (2) there exists a power product T of I_1, \dots, I_k such that $TI_{k+1} - R = D \in \langle \mathbb{C}_k \rangle$. As $\text{lv}(I_{k+1}) = y_l$, I_{k+1} does not involve y_{l+1}, \dots, y_k . This implies that $\deg(R, y_l) < \deg(I_{k+1}, y_l)$ and $\deg(R, y_i) = \deg(I_{k+1}, y_i) = 0$ for $l + 1 \leq i \leq k$. It follows that $R <_{\text{plex}} I_{k+1}$ and thus $Ry_{k+1}^d <_{\text{plex}} I_{k+1}y_{k+1}^d$, where $d = \deg(C_{k+1}, y_{k+1})$. Multiplying the two sides of $C_{k+1} = I_{k+1}y_{k+1}^d + H_{k+1}$ by T , we have

$$TC_{k+1} = TI_{k+1}y_{k+1}^d + TH_{k+1} = Ry_{k+1}^d + Dy_{k+1}^d + TH_{k+1} \in \langle \mathbb{P} \rangle.$$

Since $\deg(H_{k+1}, y_{k+1}) < d$ and T does not involve y_{k+1} , TH_{k+1} is B-reduced with respect to $\mathbb{G}^{(m+k+1)}$. Note that $D \in \langle \mathbb{C}_k \rangle \subset \langle \mathbb{G} \rangle$, so $\text{nform}(Dy_{k+1}^d, \mathbb{G}) \equiv 0$. If $R \neq 0$, then

$$0 \neq Ry_{k+1}^d + \text{nform}(TH_{k+1}, \mathbb{G}) = Ry_{k+1}^d + \text{nform}(Dy_{k+1}^d + TH_{k+1}, \mathbb{G}) \in \langle \mathbb{P} \rangle$$

is B-reduced with respect to $\mathbb{G}^{(0, \dots, m+k+1)}$, which leads to contradiction. Therefore, $\text{prem}(I_{k+1}, \mathbb{C}_k) = R \equiv 0$ and $\text{nform}(TH_{k+1}, \mathbb{G}) \equiv 0$.

Moreover, $\text{nform}(TH_{k+1}, \mathbb{G}) \equiv 0$ implies that $\text{prem}(TH_{k+1}, \mathbb{C}_k) \equiv 0$. Since \mathbb{C}_k is normal, T does not involve y_1, \dots, y_k and $\text{prem}(TI_{k+1}y_{k+1}^d, \mathbb{C}_k) = Ty_{k+1}^d \text{prem}(I_{k+1}, \mathbb{C}_k) \equiv 0$. It follows from Corollary 2.5 that

$$T \text{prem}(C_{k+1}, \mathbb{C}_k) = \text{prem}(TC_{k+1}, \mathbb{C}_k) = \text{prem}(TI_{k+1}y_{k+1}^d + TH_{k+1}, \mathbb{C}_k) \equiv 0.$$

Therefore, $\text{prem}(C_{k+1}, \mathbb{C}_k) \equiv 0$ and the proof is complete. \square

Lemma 3.7 Let $[C_1, \dots, C_r]$ be the W -characteristic set of $\langle \mathbb{P} \rangle$ with $\mathbb{C}_i = [C_1, \dots, C_i]$ for $1 \leq i \leq r$ and let k be the biggest integer such that \mathbb{C}_k is normal. Assume that $k < r$ and let $I_{k+1} = \text{ini}(C_{k+1})$, $y_l = \text{lv}(I_{k+1})$, and $Q = \text{pquo}(C_l, I_{k+1}, y_l)$. Then:

- (a) \mathbb{C}_{k+1} is not regular;
- (b) if I_{k+1} is R -reduced with respect to C_l , then

$$\text{prem}(C_l, [C_1, \dots, C_{l-1}, I_{k+1}]) \equiv 0 \quad \text{and} \quad \text{prem}(QC_{k+1}, \mathbb{C}_k) \equiv 0.$$

Proof. (a) Let $y_i = \text{lv}(C_i)$ for $1 \leq i \leq r$ and $R = \text{res}(I_{k+1}, \mathbb{C}_l)$. Then according to (3) there exist polynomials $A, B_1, \dots, B_l \in \mathcal{K}[u_1, \dots, u_m, y_1, \dots, y_l]$ such that

$$R = AI_{k+1} + B_1C_1 + \dots + B_lC_l, \quad (4)$$

where R does not involve y_1, \dots, y_l . Write $C_{k+1} = I_{k+1}y_{k+1}^d + H_{k+1}$, where $d = \deg(C_{k+1}, y_{k+1})$. Multiplying the two sides of this equality by A and using (4), one obtains

$$Ry_{k+1}^d + AH_{k+1} = AC_{k+1} + (B_1C_1 + \dots + B_lC_l)y_{k+1}^d \in \langle \mathbb{P} \rangle.$$

Suppose that \mathbb{C}_{k+1} is regular. Then $R \neq 0$. Note that R does not involve y_1, \dots, y_r , so R is R -reduced with respect to \mathbb{C}_k and B -reduced with respect to $\mathbb{G}^{(0, \dots, m+k)}$; thereby $R <_{\text{plex}} I_{k+1}$ and $Ry_{k+1}^d <_{\text{plex}} I_{k+1}y_{k+1}^d$. Hence Ry_{k+1}^d is B -reduced with respect to $\mathbb{G}^{(0, \dots, m+k+1)}$. It follows that

$$0 \neq Ry_{k+1}^d + \text{nform}(AH_{k+1}, \mathbb{G}) \in \langle \mathbb{P} \rangle$$

is B -reduced with respect to \mathbb{G} . This leads to contradiction. Therefore, $R \equiv 0$ and \mathbb{C}_{k+1} is not regular.

- (b) Suppose that I_{k+1} is R -reduced with respect to C_l and let

$$M = \text{prem}(\text{prem}(C_l, I_{k+1}, y_l), \mathbb{C}_{l-1}),$$

$I = \text{ini}(I_{k+1})$, and $I_i = \text{ini}(C_i)$ for $1 \leq i \leq k$. Then there exist an integer $q \geq 0$ and a power product U of I_1, \dots, I_{l-1} such that

$$U(I^qC_l - QI_{k+1}) - M = E \in \langle \mathbb{C}_{l-1} \rangle. \quad (5)$$

Recall that $C_{k+1} = I_{k+1}y_{k+1}^d + H_{k+1}$. It follows that

$$UQC_{k+1} = (I^qUC_l - M)y_{k+1}^d - Ey_{k+1}^d + UQH_{k+1} \in \langle \mathbb{P} \rangle.$$

As I_{k+1} is R -reduced with respect to C_l , $\deg(I_{k+1}, y_l) < \deg(C_l, y_l)$. Note that M does not involve y_{l+1}, \dots, y_k , so M is R -reduced with respect to \mathbb{C}_k and B -reduced with respect to $\mathbb{G}^{(0, \dots, m+k)}$. Moreover, M is R -reduced with respect to I_{k+1} and thus $M <_{\text{plex}} I_{k+1}$ and $My_{k+1}^d <_{\text{plex}} I_{k+1}y_{k+1}^d$. Hence My_{k+1}^d is B -reduced with respect to $\mathbb{G}^{(0, \dots, m+k+1)}$. If $M \neq 0$, then

$$\begin{aligned} 0 &\neq -My_{k+1}^d + \text{nform}(UQH_{k+1}, \mathbb{G}) \\ &= -My_{k+1}^d + \text{nform}(I^qUC_ly_{k+1}^d - Ey_{k+1}^d + UQH_{k+1}, \mathbb{G}) \in \langle \mathbb{P} \rangle \end{aligned}$$

is B -reduced with respect to \mathbb{G} , which leads to contradiction. Therefore, M must be identically equal to 0.

Since \mathbb{C}_k is normal, $I^qUC_l \in \langle \mathbb{C}_l \rangle$, and $E \in \langle \mathbb{C}_{l-1} \rangle$, we have $I^qUC_l - E \in \langle \mathbb{C}_l \rangle$ and

$$\text{prem}(UQI_{k+1}, \mathbb{C}_l) = \text{prem}(I^qUC_l - E, \mathbb{C}_l) \equiv 0.$$

It follows that $\text{prem}(UQI_{k+1}y_{k+1}^d, \mathbb{C}_k) \equiv 0$ and $UQH_{k+1} \in \langle \mathbb{P} \rangle$. As $\deg(H_{k+1}, y_{k+1}) < d$, UQH_{k+1} is R-reduced with respect to C_{k+1} ; thereby

$$\text{prem}(UQH_{k+1}, \mathbb{C}_k) = \text{prem}(UQH_{k+1}, \mathbb{C}_{k+1}) \equiv 0.$$

Hence $\text{prem}(UQC_{k+1}, \mathbb{C}_k) \equiv 0$ by Corollary 2.5. As U does not involve y_1, \dots, y_k ,

$$\text{prem}(UQC_{k+1}, \mathbb{C}_k) = U \text{prem}(QC_{k+1}, \mathbb{C}_k).$$

Therefore, $\text{prem}(QC_{k+1}, \mathbb{C}_k) \equiv 0$. The proof is complete. \square

Lemma 3.8 *Let $[C_1, \dots, C_r]$ be the W-characteristic set of $\langle \mathbb{P} \rangle$ and k be the biggest integer such that $[C_1, \dots, C_k]$ is normal. Assume that $k < r$ and let $I_{k+1} = \text{ini}(C_{k+1})$ and $y_l = \text{lv}(I_{k+1})$. If I_{k+1} is R-reduced with respect to C_l , then either $\text{res}(\text{ini}(I_{k+1}), \mathbb{C}_{l-1}) \equiv 0$, or*

$$\text{prem}(C_{k+1}, [C_1, \dots, C_{l-1}, I_{k+1}, C_{l+1}, \dots, C_k]) \equiv 0.$$

Proof. Suppose that I_{k+1} is R-reduced with respect to C_l , let $y_i = \text{lv}(C_i)$, $I_i = \text{ini}(C_i)$, and $\mathbb{C}_i = [C_1, \dots, C_i]$ for $1 \leq i \leq r$, and let

$$\tilde{\mathbb{C}} = [C_1, \dots, C_{l-1}, I_{k+1}, C_{l+1}, \dots, C_k], \quad d = \deg(C_{k+1}, y_{k+1}), \quad H_{k+1} = C_{k+1} - I_{k+1}y_{k+1}^d,$$

and $R = \text{prem}(H_{k+1}, \tilde{\mathbb{C}})$. Then there exist nonnegative integers $s, s_1, \dots, s_{l-1}, s_{l+1}, \dots, s_k$ and polynomials $B, B_1, \dots, B_{l-1}, B_{l+1}, \dots, B_k \in \mathcal{K}[u_1, \dots, u_m, y_1, \dots, y_k]$ such that

$$I^s I_1^{s_1} \dots I_{l-1}^{s_{l-1}} I_{l+1}^{s_{l+1}} \dots I_k^{s_k} QH_{k+1} = BQI_{k+1} - QB_l C_l + \sum_{i=1}^k QB_i C_i + QR, \quad (6)$$

where $Q = \text{pquo}(C_l, I_{k+1}, y_l)$ and $I = \text{ini}(I_{k+1})$. The first conclusion of Lemma 3.7 (b) implies that $L = \text{prem}(C_l, I_{k+1}, y_l) \in \text{sat}(\mathbb{C}_{l-1})$, i.e., $I^q C_l - QI_{k+1} = L \in \text{sat}(\mathbb{C}_{l-1})$ for some integer $q \geq 0$. Recall that \mathbb{C}_k is normal. Hence $QI_{k+1} \in \text{sat}(\mathbb{C}_l)$ and $\text{prem}(QI_{k+1}, \mathbb{C}_l) \equiv 0$. This, together with the second conclusion of Lemma 3.7 (b), implies that $\text{prem}(QH_{k+1}, \mathbb{C}_k) = \text{prem}(QC_{k+1} - QI_{k+1}y_{k+1}^d, \mathbb{C}_k) \equiv 0$, so that $QH_{k+1} \in \text{sat}(\mathbb{C}_k)$. It follows from (6) that $QR \in \text{sat}(\mathbb{C}_k)$.

As I_{k+1} is R-reduced with respect to C_l , $\deg(Q, y_l) > 0$. From the pseudo-remainder formula $I^q C_l = QI_{k+1} + L$, one sees that $\text{ini}(Q) = I^{q-1} I_l$ for some $q \geq 0$. Note that I_l does not involve y_1, \dots, y_{l-1} and assume that $\text{res}(I, \mathbb{C}_{l-1}) \not\equiv 0$. Then $M = \text{res}(I^{q-1} I_l, \mathbb{C}_{l-1}) \not\equiv 0$. Hence there exists a polynomial $S \in \mathcal{K}[u_1, \dots, u_m, y_1, \dots, y_{l-1}]$ such that $M - S \text{ini}(Q) = A \in \langle \mathbb{C}_{l-1} \rangle$. Write $Q = \text{ini}(Q)y_l^\delta + \bar{Q}$ and let $Z = My_l^\delta + S\bar{Q}$, where $\delta = \deg(Q, y_l)$. Then $SQ = Z - Ay_l^\delta$ and thus $SQR = ZR - ARy_l^\delta$. It follows that $ZR \in \text{sat}(\mathbb{C}_k)$. Since \mathbb{C}_k is normal, $\text{prem}(ZR, \mathbb{C}_k) \equiv 0$ according to Proposition 2.3. It is easy to see that

$$\begin{aligned} \deg(ZR, y_l) &= \deg(Z, y_l) + \deg(R, y_l) = \delta + \deg(R, y_l) \\ &= \deg(C_l, y_l) - \deg(I_{k+1}, y_l) + \deg(R, y_l) < \deg(C_l, y_l). \end{aligned}$$

Hence

$$\text{prem}(MRy_l^\delta + S\bar{Q}R, \mathbb{C}_{l-1}) = \text{prem}(ZR, \mathbb{C}_{l-1}) = \text{prem}(ZR, \mathbb{C}_l) = \text{prem}(ZR, \mathbb{C}_k) \equiv 0$$

(for Z does not involve y_{l+1}, \dots, y_k). This implies that $MR = \text{prem}(MR, \mathbb{C}_{l-1}) \equiv 0$ according to Lemma 2.4. Therefore, $R \equiv 0$.

As \mathbb{C}_k is normal and $\text{res}(I, \mathbb{C}_{l-1}) \not\equiv 0$, $\tilde{\mathbb{C}}$ is regular. Obviously, $\text{prem}(I_{k+1}y_{k+1}^d, \tilde{\mathbb{C}}) = y_{k+1}^d \text{prem}(I_{k+1}, \tilde{\mathbb{C}}) \equiv 0$. Hence

$$\text{prem}(C_{k+1}, \tilde{\mathbb{C}}) = \text{prem}(I_{k+1}y_{k+1}^d + H_{k+1}, \tilde{\mathbb{C}}) \equiv 0$$

by Corollary 2.5. The proof is complete. \square

Theorem 3.9 Let $\mathbb{C} = [C_1, \dots, C_r]$ be the W -characteristic set of $\langle \mathbb{P} \rangle$ and $\mathbb{C}_i = [C_1, \dots, C_i]$ for $1 \leq i \leq r$. If \mathbb{C} is abnormal, then there exists an integer k ($1 \leq k < r$) such that

- (a) \mathbb{C}_k is normal and thus regular;
- (b) \mathbb{C}_{k+1} is not regular;
- (c) if $I_{k+1} = \text{ini}(C_{k+1})$ is not R -reduced with respect to C_l , then

$$\text{prem}(I_{k+1}, \mathbb{C}_l) \equiv 0 \quad \text{and} \quad \text{prem}(C_{k+1}, \mathbb{C}_k) \equiv 0;$$

- (d) if I_{k+1} is R -reduced with respect to C_l , then $\text{prem}(C_l, [C_1, \dots, C_{l-1}, I_{k+1}]) \equiv 0$ and either $\text{res}(\text{ini}(I_{k+1}), \mathbb{C}_{l-1}) \equiv 0$, or

$$\text{prem}(C_{k+1}, [C_1, \dots, C_{l-1}, I_{k+1}, C_{l+1}, \dots, C_k]) \equiv 0,$$

where $y_l = \text{lv}(I_{k+1})$.

Proof. Suppose that \mathbb{C} is abnormal and let k ($1 \leq k < r$) be the biggest integer such that \mathbb{C}_k is normal. Then we have (a). By Lemma 3.7 (a), \mathbb{C}_{k+1} is not regular; thus (b) is proved. The identity $\text{prem}(C_{k+1}, \mathbb{C}_k) \equiv 0$ in (c) has been proved as the second conclusion of Theorem 3.6. Recall $\text{prem}(I_{k+1}, \mathbb{C}_k) \equiv 0$, the first conclusion of Theorem 3.6, where $I_{k+1} = \text{ini}(C_{k+1})$. As y_{l+1}, \dots, y_k do not appear in I_{k+1} , one sees that $\text{prem}(I_{k+1}, \mathbb{C}_k) = \text{prem}(I_{k+1}, \mathbb{C}_l)$. Hence $\text{prem}(I_{k+1}, \mathbb{C}_l) \equiv 0$ and (c) is proved.

The identity $\text{prem}(C_l, [C_1, \dots, C_{l-1}, I_{k+1}]) \equiv 0$ in (d) has been proved as the first conclusion of Lemma 3.7 (b), and so has the second conclusion of (d) proved as Lemma 3.8. \square

The irregularity index $m + k + 1$ in Theorem 3.9 is clearly characteristic. The W -characteristic set of the $(n - m - k - 1)$ th elimination ideal $\langle \mathbb{P} \rangle^{(0, \dots, m+k+1)}$ is irregular, while that of the $(n - m - k)$ th elimination ideal $\langle \mathbb{P} \rangle^{(0, \dots, m+k)}$ is not only regular but also normal. As shown in Corollary 3.5, from the normal W -characteristic set of the elimination ideal a Ritt characteristic set of the ideal can be computed rather easily by means of pseudo-division. When the W -characteristic set of the ideal $\langle \mathbb{P} \rangle$ is itself normal, $m + k + 1$ may be defined to be $n + 1$ (or any other integer greater than n) and the Buchberger-Gröbner basis \mathbb{G} of \mathbb{P} may be said to be *regular*.

The pseudo-divisibility relations in Theorem 3.9 (c) and (d) expose the intrinsic structure of the polynomials C_l and C_{k+1} and the irregularity of C_{k+1} modulo the saturated ideal of the normal triangular set \mathbb{C}_k . More relations of this kind would help us gain more insights into the structure of the polynomials in \mathbb{G} .

3.3 Examples

The following examples serve to illustrate various behaviors of the W -characteristic sets of polynomial ideals.

Example 3.1 (a) Let $\mathbb{P} = \{x_1x_2 - 1, x_3 - x_2\}$. Then the W -characteristic set of $\langle \mathbb{P} \rangle$ is $\mathbb{C} = [x_1x_2 - 1, x_3 - x_2]$. \mathbb{C} is normal, but it is not a Ritt characteristic set of $\langle \mathbb{P} \rangle$. Construct $\mathbb{C}^* = [x_1x_2 - 1, \text{prem}(x_3 - x_2, [x_1x_2 - 1])] = [x_1x_2 - 1, x_1x_3 - 1]$. Then \mathbb{C}^* is a Ritt characteristic set of $\langle \mathbb{P} \rangle$ and $\mathbb{C}^* \sim \mathbb{C}$, but $x_1x_3 - 1 >_{\text{plex}} x_3 - x_2$.

(b) Let $\mathbb{P} = \{x_1^2, (x_2 + x_1)x_3 + x_1\}$. Then the W-characteristic set of $\langle \mathbb{P} \rangle$ is $\mathbb{C} = [x_1^2, (x_2 + x_1)x_3 + x_1]$: it is a regular ascending set and thus is a Ritt characteristic set of $\langle \mathbb{P} \rangle$. The regular set \mathbb{C} is not normal because the parameter x_2 is ordered greater than the leading variable x_1 . This example explains why the assumption on the variable order is necessary for the W-characteristic set of $\langle \mathbb{P} \rangle$ to be normal or exhibit irregularity structure.

(c) Let $\mathbb{P} = \{x_1x_2, x_2x_3, x_3x_4\}$. Then the W-characteristic set of $\langle \mathbb{P} \rangle$ is $\mathbb{C} = [C_1, C_2, C_3] = [x_1x_2, x_2x_3, x_3x_4]$. \mathbb{C} is abnormal and irregular and it is not a Ritt characteristic set of $\langle \mathbb{P} \rangle$. One may see that $\text{prem}(C_2, C_1, x_2) = \text{prem}(C_3, C_2, x_3) \equiv 0$ and $[x_1x_2, x_3x_4]$ is a Ritt characteristic set of $\langle \mathbb{P} \rangle$. However, application of the construction for C_i^* in Theorem 3.4 to the irregular W-characteristic set \mathbb{C} here does not lead to the Ritt characteristic set $[x_1x_2, x_3x_4]$ of $\langle \mathbb{P} \rangle$.

(d) Let $C_3 = x_1(x_2 + x_1)x_3 - x_1^3$ and $\mathbb{P} = \{x_1^4, x_2^4, C_3\}$. Then $\mathbb{G} = \{x_1^4, x_1^3x_2^3, x_2^4, C_3\}$. Thus the W-characteristic set of $\langle \mathbb{P} \rangle$ is $\mathbb{C} = [x_1^4, x_1^3x_2^3, C_3]$, which is abnormal and irregular. By Theorem 3.3, \mathbb{C} is a Ritt characteristic set of $\langle \mathbb{P} \rangle$.

Now let $C_1 = x_1^3$, $C_2 = x_2^3$, $\bar{C}_3 = x_1x_2x_3 - x_1^2x_2$ and $\bar{\mathbb{P}} = \{C_1, C_2, \bar{C}_3\}$. Then the W-characteristic set $\bar{\mathbb{C}} = [C_1, C_2, \bar{C}_3]$ of $\langle \bar{\mathbb{P}} \rangle$ is a Ritt characteristic set of $\langle \bar{\mathbb{P}} \rangle$ by Theorem 3.3. One can see that $Q = \text{pquo}(C_2, \text{ini}(\bar{C}_3), x_2) = x_1^2x_2^2$ and $\text{res}(\text{ini}(Q), [C_1]) \equiv 0$. This example shows that the case in which $\text{res}(I, \mathbb{C}_{l-1}) \equiv 0$ in Theorem 3.9 (d) does occur.

(e) Let $C_1 = x_1x_2$, $C_2 = x_3x_4 - x_2^2$, $C_3 = x_2x_5 + x_4^2$ and $\mathbb{P} = \{C_1, C_2, C_3\}$. Then $\mathbb{G} = \{C_1, C_2, x_1x_4^2, C_3\}$ and the W-characteristic set of $\langle \mathbb{P} \rangle$ is $\mathbb{C} = [C_1, C_2, C_3]$. One can verify that $\text{prem}(C_3, [C_1, C_2]) \equiv 0$ and $\text{prem}(x_1x_5 - x_1, \mathbb{C}) \equiv 0$, but $x_1x_5 - x_1 \notin \langle \mathbb{P} \rangle$.

Let $\bar{C}_2 = x_2x_4 - x_2^2$ instead of C_2 and $\bar{\mathbb{P}} = \{C_1, \bar{C}_2, C_3\}$. Then $\bar{\mathbb{G}} = \{C_1, \bar{C}_2, x_1x_4^2, x_4^3 - x_2^3, C_3\}$ and the W-characteristic set of $\langle \bar{\mathbb{P}} \rangle$ is $\bar{\mathbb{C}} = [C_1, \bar{C}_2, C_3]$. Now $\text{prem}(C_3, [C_1, \bar{C}_2]) = \text{prem}(\bar{C}_2, [C_1]) \equiv 0$. However, $x_1x_4^2 \in \langle \bar{\mathbb{P}} \rangle$, but $\text{prem}(x_1x_4^2, [C_1]) \neq 0$. Therefore, $[C_1]$ is not a Ritt characteristic set of $\langle \bar{\mathbb{P}} \rangle$, and we suspect that $[C_1, x_1x_4^2]$ is. This example shows that, in the abnormal case, the minimal ascending set contained in the W-characteristic set \mathbb{C} of an ideal \mathcal{I} is not necessarily a Ritt characteristic set of the ideal and the Buchberger-Gröbner basis \mathbb{G} of \mathcal{I} may contain other ascending sets of lower rank. A natural question that remains to be answered is how to construct a Ritt characteristic set of \mathcal{I} from \mathbb{G} when \mathbb{C} is neither an ascending set nor a regular set.

For Example 3.1 (c), as well as other examples we have studied, it appears that the Ritt characteristic set of an ideal \mathcal{I} does not characterize the ideal well enough in the abnormal case. This is caused essentially by the irregularity of the ideal. The W-characteristic set of \mathcal{I} , which provides sufficient information about \mathcal{I} , may serve as an alternative to the Ritt characteristic set.

4 Some Remarks

We point out two directions of research in which triangular sets and Buchberger-Gröbner bases have been explored reciprocally. The first is concerned with algorithmic decomposition of polynomial or differential polynomial sets (or systems) into triangular or differential triangular sets (or systems) of various kinds, where the method of Buchberger-Gröbner bases is used as a black-box tool to handle some of the involved algebraic computational issues. The second direction is devoted to the investigation of alternative algorithms from the constructive theory of partial differential equations developed by C. H. Riquier, M. Janet, J. F. Ritt, J. M. Thomas, and others for efficient computation of Gröbner bases with variants. The literature is rich for each of these directions and there is a large amount of work which

may be considered as relevant to what is presented here. A review of such related work is beyond the intended scope of this paper.

The connection we have established between Ritt's characteristic sets and Buchberger's Gröbner bases is expected to stimulate further research and development on some of the outstanding problems in the above-mentioned directions. For example, the pseudo-divisibility relations shown in Theorem 3.9 (c) and (d) allow us to split the Buchberger-Gröbner basis \mathbb{G} by using the explicit and nontrivial factorizations of C_l and C_{k+1} to compute a normal triangular decomposition. Let us explain (part of) the splitting process briefly.

1. When I_{k+1} is not R-reduced with respect to C_l , the first conclusion of Theorem 3.9 (c) implies that $I_1^{t_1} \cdots I_l^{t_l} I_{k+1} \in \langle C_l \rangle \subset \langle \mathbb{G} \rangle$ for some integers $t_i \geq 0$, where $I_i = \text{ini}(C_i)$ for $1 \leq i \leq r$. Note that I_1, \dots, I_{k+1} are all B-reduced with respect to \mathbb{G} . Now \mathbb{G} can be split into $\mathbb{G} \cup \{I_1\}, \dots, \mathbb{G} \cup \{I_l\}, \mathbb{G} \cup \{I_{k+1}\}$.

2. To deal with the case in which I_{k+1} is R-reduced with respect to C_l , we recall the pseudo-remainder formula $I^q C_l = QI_{k+1} + L$ for $\text{prem}(C_l, I_{k+1}, y_l) = L$, where $I = \text{ini}(I_{k+1})$ and $q \geq 0$. One can easily see that $\text{ini}(Q) = I^{q-1}I_l$. Suppose that $\text{prem}(I^{q-1}I_l, C_{l-1}) = I_l \text{prem}(I^{q-1}, C_{l-1}) \equiv 0$, so $\text{prem}(I^{q-1}, C_{l-1}) \equiv 0$. Then according to (2) there exist nonnegative integers s_1, \dots, s_{l-1} such that $I_1^{s_1} \cdots I_{l-1}^{s_{l-1}} I^{q-1} \in \langle C_{l-1} \rangle \subset \langle \mathbb{G} \rangle$. Of course, I is B-reduced with respect to \mathbb{G} . Now \mathbb{G} can be split into $\mathbb{G} \cup \{I_1\}, \dots, \mathbb{G} \cup \{I_{l-1}\}, \mathbb{G} \cup \{I\}$.

3. According to the first conclusion of Theorem 3.9 (d), $\text{prem}(L, C_{l-1}) \equiv 0$. This implies that $I^q C_l - QI_{k+1} \in \text{sat}(C_{l-1})$ and thus $QI_{k+1} \in \text{sat}(C_l)$. Therefore, $I_1^{q_1} \cdots I_{l-1}^{q_{l-1}} QI_{k+1} \in \langle C_l \rangle \subset \langle \mathbb{G} \rangle$ for some integers $q_i \geq 0$. As $\deg(I_{k+1}, y_l) > 0$, Q is obviously R-reduced with respect to C_l . Now suppose that $\text{prem}(\text{ini}(Q), C_{l-1}) = \text{prem}(I^{q-1}I_l, C_{l-1}) \neq 0$. It follows from Lemma 2.4 that $\text{prem}(Q, C_{l-1}) \neq 0$. Clearly, $\text{prem}(Q, C_{l-1})$ is R-reduced with respect to C_l and thus B-reduced with respect to \mathbb{G} . Then \mathbb{G} can be split into $\mathbb{G} \cup \{I_1\}, \dots, \mathbb{G} \cup \{I_l\}, \mathbb{G} \cup \{\text{prem}(Q, C_{l-1})\}, \mathbb{G} \cup \{I_{k+1}\}$.

In any case of splitting, the split polynomial set \mathbb{G}^+ is obtained from \mathbb{G} by adjoining a nonzero polynomial $F \in \mathcal{K}[u_1, \dots, u_m, y_1, \dots, y_l]$ which is B-reduced with respect to \mathbb{G} and thus does not belong to $\langle \mathbb{G} \rangle$. Hence $\langle \mathbb{G} \rangle \subset \langle \mathbb{G}^+ \rangle$ and $\langle \mathbb{G} \rangle \neq \langle \mathbb{G}^+ \rangle$.

Consider further the reduced Buchberger-Gröbner basis and the W-characteristic set of each $\langle \mathbb{G}^+ \rangle$ and continue the splitting process. In view of the Ascending Chain Condition of polynomial ideals [4, Ch. 2, Theorem 7], the splitting process must terminate in a finite number of steps. Finally, we shall obtain finitely many reduced plex Buchberger-Gröbner bases $\mathbb{G}_1, \dots, \mathbb{G}_e$ such that $\text{Zero}(\mathbb{P}) = \text{Zero}(\mathbb{G}_1) \cup \dots \cup \text{Zero}(\mathbb{G}_e)$, or equivalently $\sqrt{\langle \mathbb{P} \rangle} = \sqrt{\langle \mathbb{G}_1 \rangle} \cap \dots \cap \sqrt{\langle \mathbb{G}_e \rangle}$, and the W-characteristic set \mathbb{C}_i of each $\langle \mathbb{G}_i \rangle$ is normal for $1 \leq i \leq e$. In the case when $\text{sat}(\mathbb{C}_j) \neq \langle \mathbb{G}_j \rangle$ which can be determined, for instance, by computing the reduced plex Buchberger-Gröbner basis \mathbb{G}'_j of $\text{sat}(\mathbb{C}_j)$, \mathbb{G}_j may be further split into \mathbb{G}'_j (called a *strong* regular Buchberger-Gröbner basis) and $\mathbb{G}_j \cup \{F_j\}$, where F_j is the product of the initials of the polynomials in \mathbb{C}_j . Therefore, we may also ensure that $\text{sat}(\mathbb{C}_i) = \langle \mathbb{G}_i \rangle$ for all i ($1 \leq i \leq e$).

The process described above can be formulated as an algorithm. We shall detail the algorithm and discuss computational aspects elsewhere.

We have no idea how to extend the presented connection from the algebraic to the differential case because a general theory of Gröbner bases for differential polynomial ideals is still lacking. This paper has benefited from the discussions which the author had with Xiaoliang Li, Chenqi Mou, and Jing Yang.

References

- [1] Becker, T., Weispfenning, V. (1993): Gröbner bases: a computational approach to commutative algebra. Springer, New York Berlin Heidelberg.
- [2] Buchberger, B. (1965): Ein Algorithmus zum Auffinden der Basiselemente des Restklassenringes nach einem nulldimensionalen Polynomideal. Ph.D. thesis, Universität Innsbruck, Austria.
- [3] Buchberger, B. (1985): Gröbner bases: an algorithmic method in polynomial ideal theory. In: Bose, N. K. (ed.): Multidimensional systems theory. Reidel, Dordrecht, pp. 184–232.
- [4] Cox, D., Little, J., O’Shea, D. (1996): Ideals, varieties, and algorithms (2nd edn.). Springer, New York Berlin Heidelberg.
- [5] Eisenbud, D. (1995): Commutative algebra: with a view toward algebraic geometry. Springer, New York.
- [6] Miller, E., Sturmfels, B. (2005): Combinatorial commutative algebra. Springer, New York.
- [7] Mishra, B. (1993): Algorithmic algebra. Springer, New York.
- [8] Ritt, J. F. (1932): Differential equations from the algebraic standpoint. American Mathematical Society, New York.
- [9] Ritt, J. F. (1950): Differential algebra. American Mathematical Society, New York.
- [10] Wang, D. (2001): Elimination methods. Springer, Wien New York.
- [11] Wu, W.-t. (1994): Mechanical theorem proving in geometries: basic principles. Springer, Wien New York [translated from the Chinese edition — published in 1984 by Science Press, Beijing — by X. Jin and D. Wang].